IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received April 20, 2021, accepted May 25, 2021, date of publication June 3, 2021, date of current version June 23, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3085896

Comparing Wi-Fi 6 and 5G Downlink Performance for Industrial IoT

ROBERTO MALDONADO[1], ANDERS KARSTENSEN[2], GUILLERMO POCOVI[1], ALI A. ESSWIE[1], CLAUDIO ROSA[1], OLLI ALANEN[3], MIKA KASSLIN[3], (Member, IEEE), AND TROELS KOLDING[1]

[1]Nokia Bell Labs, 9220 Aalborg, Denmark
[2]Department of Electronic Systems, Aalborg University, 9220 Aalborg, Denmark
[3]Nokia Bell Labs, 02610 Espoo, Finland

Corresponding author: Troels Kolding (troels.kolding@nokia-bell-labs.com)

ABSTRACT The use of wireless communications in Industrial Internet of Things (IIoT) enables unparalleled levels of flexibility and instantaneous reconfiguration for autonomous industrial processes. In this paper, the focus is on optimizing and evaluating Wi-Fi 6 and 5G New Radio (NR) licensed and unlicensed wireless networks for meeting the packet latency and reliability requirements of critical IIoT applications. The study is based on extensive system simulations using a 3GPP-defined IIoT indoor factory framework and application traffic models. Each radio technology is individually optimized leveraging the pros and cons of that technology to maximize the carried load in the network while fulfilling the delay requirements at a specified reliability level of 99.999 %. In addition to a performance comparison, the paper also provides deployment guidance for applying each radio technology in the considered IIoT setting. With proposed latency aware scheduling and when operated in interference free spectrum, Wi-Fi 6 can support <1 ms applications at a very low load, whereas the performance gap with respect to 5G NR reduces as delay requirements are relaxed to 10-100 ms. Conditioned on the fulfilment of the application latency and reliability requirements, unlicensed 5G NR shows nearly $2\times$ the spectral efficiency of Wi-Fi 6 in all available configurations. Licensed 5G NR shows generally the best performance, especially for delay requirement <1 ms, supporting 2-4$\times$ the spectral efficiency achievable by unlicensed technologies.

INDEX TERMS Industrial wireless communications, 5G new radio licensed and unlicensed, Wi-Fi 6.

I. INTRODUCTION

One enabler for Industry 4.0 is pervasive connectivity between machines, people, and objects. Wireless communications adds value to industrial communications in several ways, including reliable connectivity to moving objects such as mobile robots, automated guided vehicles (AGVs), drones, and humans as well as removing cables from stationary, rotating, or other objects with limited mobility. These exciting new prospects come with open questions for vertical industries. Which wireless technology meets the use-case requirements? Is licensed spectrum needed or is unlicensed spectrum sufficient? The objective of the present research is to clarify some of those questions when targeting application requirements found in Industrial IoT (IIoT). 5G New Radio (NR), including unlicensed (NR-U) and licensed Frequency Division Duplex and Time Division Duplex variants (NR-FDD, NR-TDD), feature Ultra Reliable Low Latency Communications (URLLC) supporting sub-ms air interface delay with up to 99.9999 % reliability. Hence, 5G is designed for latency critical applications in the IIoT domain [1], [2]. Wi-Fi 6 is an often considered alternative as Wi-Fi is already present in many IIoT environments for non-critical applications and it comes with significant performance improvements with the latest IEEE 802.11ax extension (Wi-Fi 6).

While Wi-Fi and 5G NR latency performance for critical applications has been investigated separately in several studies [3]–[8], a comparison focused on the reliability levels expected in IIoT is not available. A general comparison of NR-U and Wi-Fi using average metrics was conducted in [9]. Our contribution is a comprehensive comparison of

The associate editor coordinating the review of this manuscript and approving it for publication was Jose Saldana.

the downlink radio performance of Wi-Fi 6, NR-U, NR-FDD, and NR-TDD in a well-defined indoor factory IIOT scenario, considering various requirements ranging from strictly delay critical applications (<1 ms) such as motion control to more delay tolerant applications (<100 ms) such as Manufacturing Execution System (MES) applications. As focus is on critical IIoT, we focus on reliability, meaning that, for the wireless deployment to be accepted, latency guarantees have to be met with a 99.999% probability.

Wi-Fi 6 and 5G are very different radio technologies in terms of their ability to re-use the same spectrum among multiple base stations and to which extent network operations are synchronized. Hence, Wi-Fi 6 and 5G planners must use a very different mindset when optimizing the deployment for IIoT use-cases. A main contribution of the presented research is providing a fair comparison where each technology is optimized to its strengths with the target of maximizing the capacity for the IIoT environment while ensuring application requirements are strictly met. This includes aligned and optimized IIoT radio resource management algorithms tailored to the specific radio technologies, of interest to Wi-Fi 6 and 5G vendors and planners.

The paper is organized as follows. First, in Section II, the considered indoor factory setting is detailed as well as deployment and bandwidth assumptions. This section includes also basic information related to licensed and unlicensed spectrum as well as Wi-Fi 6, 5G NR, and 5G NR-U technologies. In Section III individual optimizations of Wi-Fi 6, 5G NR-U, and 5G NR are presented including results showing the impacts of key deployment decisions. The overall performance comparison between the technologies is presented in Section IV and conclusions are drawn in Section V.

II. SYSTEM MODEL

A. IIoT ENVIRONMENT

As reference scenario for wireless IIoT, we consider the Indoor-Factory sparse-clutter (InF-SH) model defined by the 3rd Generation Partnership Project (3GPP) [10]. As shown in Fig. 1, the scenario resembles an indoor factory hall of 120 × 50 × 5 cubic meters where 20 % of the total surface is occupied by clutter with an average height of 2 m. Base stations (BS, i.e. 5G NR/NR-U gNBs and Wi-Fi access points APs) are mounted in the ceiling in a location that depends on the number of installed gNBs/APs. Deployment scenarios with 1, 2, 4, 8 and 12 BS in the factory hall are considered as part of the network optimizations specific of the different radio access technologies (RAT). For the deployment of devices, we consider 60 static devices in total in the factory hall, spatially uniformly distributed over the factory area. 5G devices are denoted user equipment (UE) while Wi-Fi devices are denoted as stations (STA). It is assumed that the deployment is carefully planned to achieve even load, so that a fixed number of connected devices per base station

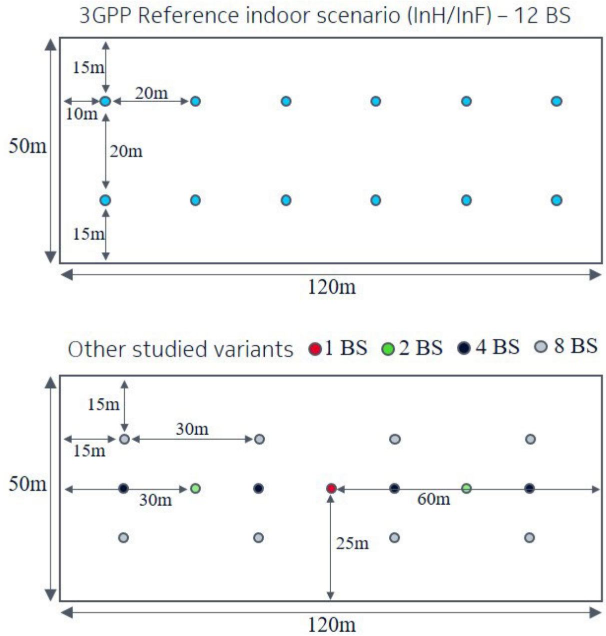


FIGURE 1. Indoor factory scenario including base station installation points.

is assumed, e.g. if there are 2 BS deployed, there will be 30 devices per BS.

While cable-like reliability is in focus in wireless IIoT, the application delay requirements vary significantly depending on the use-case. Motion control applications may have 1-4 ms delay requirements, while alarm and control mechanisms may tolerate 10 ms delay. The different traffic classes considered to be representative of different use-cases are shown in Table 1 and have been chosen to see more consistent trends when scaling the delay requirement. All traffic classes are based on the FTP Model 3 from [11], where packet arrival is modeled as a Poisson Process, with average arrival rate λ packets per second, and a packet size (B) of 50 B. Note that periodic traffic (i.e. fixed inter-arrival time) is also a common assumption for industrial applications, especially motion control; however, this type of traffic is not considered in this work due to the lack of modelling assumptions for the relative offset of the timing of traffic arrivals across UEs which play a major role in the obtained performance. We refer to [8] for 5G NR performance results with periodic traffic and random (uniformly-distributed) arrival offset. Performance results are presented in Section III and onward for different offered load conditions in the network, where the offered load is calculated as $N \cdot B \cdot 8 \cdot \lambda$ [b/s] (where $N = 60$ is the number of deployed UEs, and 8 corresponds to the conversion from Bytes to bits)

B. SPECTRUM ACCESS

Licensed spectrum can be assigned to be used by vertical industries on an exclusive basis in specific geographical regions or deployed in collaboration with a cellular service provider. As a key quality towards critical IIoT, licensed

TABLE 1. Considered IIoT reference applications, example use-cases could be motion control (I), collaborative AGVs or robotic control (II and III), and asset/process monitoring or autonomous AGVs (IV).

Application	Packet size	Arrival model	Maximum air interface packet delay	Minimum reliability
I	50 B	Poisson	1 ms	99.999 %
II	50 B	Poisson	4 ms	99.999 %
III	50 B	Poisson	10 ms	99.999 %
IV	50 B	Poisson	100 ms	99.999 %

TABLE 2. Key parameters of the considered radio technologies.

Parameter	Wi-Fi 6	5G NR-U	5G NR-TDD	5G NR-FDD
Carrier	5 GHz	5 GHz	3.5 GHz	3.5 GHz
MIMO config.	2x2	2x2	2x2	2x2
Ant. pattern	Omni	Omni	Omni	Omni
BS TX power	23 dBm	23 dBm	27 dBm	27 dBm
Duplex scheme	TDD	TDD	TDD	FDD
DL bandwidth	40 MHz	40 MHz	40 MHz	20 MHz
SCS	78.125 kHz	30 kHz	30 kHz	30 kHz
TDD config.	Dynamic	DDU	DDU	N/A
TX interval	Dynamic	0.143 ms	0.143 ms	0.143 ms
Max. mod.	1024QAM	256QAM	256QAM	256QAM
Receiver type	IRC	IRC	IRC	IRC

spectrum can be accessed at any time and without any particular restriction. On the other hand, unlicensed spectrum is available for non-exclusive usage, subject to specific regulatory constraints that aim at guaranteeing fair access to the radio access technologies (RAT) operating on the same spectrum. Among the requirements specified in [12], the channel access mechanisms impact both latency and reliability of transmission.

There are two types of channel access available for operation on the 5 GHz unlicensed band: Load Based Equipment (LBE) and Frame Based Equipment (FBE). With LBE, an initiating device acquires the right to transmit on a channel for a restricted amount of time by performing an asynchronized clear channel assessment procedure with randomized duration and exponential back-off, in case of collisions. FBE, on the other hand, is a synchronized and deterministic channel access mechanism where channel sensing is performed based on a periodic and well-determined frame structure. FBE is not suited for good co-existence in uncontrolled environments, i.e. in presence of nearby systems and/or RATs operating on the same frequency. On the other hand, FBE is an attractive channel access mode in controlled environments, where interference is exclusively caused by devices of the same network. Further details about both channel access schemes are found in [12], [13].

C. RADIO ENVIRONMENT

The IIoT environment detailed in Section II-A includes a general radio propagation model applicable to different frequency spectrum, including 5G NR-U and Wi-Fi 6 in 5 GHz, and 5G NR in 3.5 GHz. In order to align the comparison between the technologies, we select a common channel bandwidth of 40 MHz. The choice of 40 MHz is motivated mainly by Wi-Fi considerations as discussed in Sub-section III-B, but it allows for results to be scalable to larger bandwidths. Scalability is relevant as large spectrum bandwidth for Wi-Fi and NR-U can often be assumed to be available in the 5 or 6 GHz unlicensed spectrum range. For NR-FDD, where UL and DL transmissions occur on separate frequency bands, note that only 20 MHz bandwidth (out of the 40 MHz total spectrum) is assumed to be available for DL transmission. Other basic characteristics of the radio technologies are summarized in Table 2. A 2 × 2 MIMO setup with omni-directional patterns has been assumed for all radio technologies, though more advanced base station capabilities including directional antennas and beam-forming are possible.

As commonly assumed for IIoT deployments, factories are shielded environments in which interference is tightly controlled by network planning and perimeter control. This ensures that no other RAT is deployed on the same frequency band. Under these conditions, the experienced interference is only caused by devices of the same network deployment. Controlled interference conditions are essential for supporting critical applications in the unlicensed spectrum. In scenarios with multiple RATs coexisting on the same frequency band, i.e. in uncontrolled interference conditions, the time dedicated to performing the channel access considerably limits the supported delay-sensitive applications.

The details of the optimization of each radio technology for the selected IIoT environment, including the key performance indicators considered in the optimization process, are given in the following sections.

D. EVALUATION METHODOLOGY

The evaluation methodology is based on advance dynamic system-level simulations to obtain results with a high degree of realism. The in-house developed system-level simulator uses the Monte Carlo method [14] and is built over complex mathematical models that mimic the stochastic and complex nature of mobile networks. It is designed to model multi-cell multi-user deployments, under advanced channel propagation conditions, and includes standards compliant implementations for 5G NR-U, 5G NR-TDD, 5G-FDD, and Wi-Fi 6 PHY and MAC layers. All evaluations are conducted via extensive system simulations with randomized drops of users and extended run-time to achieve sufficient accuracy at the 99.999 % packet reliability level. The simulator includes detailed modeling of the unlicensed channel access mechanisms. All random processes including location of users, traffic patterns, and radio channel modeling are common between the radio technologies ensuring a high consistency in the comparison. The simulator's implementation, including the scenario and channel model, has been calibrated with other 3GPP partners, see [15], and more details are available in [16]. In this paper, we focus on downlink data transmissions but include explicit modeling of uplink control transmissions. Uplink messages modeled include transmission of block acknowledgments for Wi-Fi 6, and channel state information (CSI) and hybrid automatic repeat request (ARQ) feedback information for NR and NR-U technologies.

III. RADIO OPTIMIZATION FOR IIoT

The optimization objective pursued in this section is to maximize the offered load that can be carried assuming one of the application types in Table 1 while strictly fulfilling the corresponding latency and reliability requirements.

A. 5G NR FDD AND TDD

5G NR operates on licensed spectrum which guarantees exclusive access to the radio channel without the need for channel-sensing mechanisms. For the selected 3.5 GHz spectrum, only unpaired spectrum is generally available meaning that only TDD operation is possible. Nevertheless, to make the evaluation as comprehensive as possible, both TDD and FDD schemes are evaluated to show the impact of the duplexing method on the achievable system capacity. We consider 5G Release-16 as the baseline for the studies.

The 5G NR comprises a large set of functionalities and features to achieve data transmissions with low latency and high reliability, see e.g. [1], [2]. As key features to reduce packet latency, a short transmission time interval (TTI) and fast UE and gNB processing times are adopted. Particularly, we assume a short TTI duration of 4 OFDM symbols (also known as mini-slot in 3GPP), i.e. 0.143 ms for 30 kHz subcarrier spacing (SCS), and *UE processing capability 2* [17] which provides a best-case (FDD, no queuing delay) latency of 0.5 ms for a single transmission, and ∼1 ms after one HARQ retransmission [18]. The queuing delay accounts for the time the data arrives at the serving cell until it is considered for scheduling (transmission). Naturally, the queuing delay increases with the offered load, and these effects are captured in the presented performance results.

Second, dynamic link adaptation is employed by adjusting the modulation and coding scheme (MCS) for each downlink transmission according to the channel Quality Indicator (CQI) reports from the UE. At low offered load, the gNB selects the MCS targeting a relatively-low Block Error Rate (BLER) target of 1 %. This reduces the need for time-consuming HARQ retransmissions, which translates into a latency improvement. However, as the load increases, the BLER target is gradually increased from 1 % to 10 %. This more aggressive link adaptation is beneficial for the overall latency performance, since the reduction of the queuing and scheduling delay (from using higher modulation and code rate) compensates for the delay introduced by the increased number of HARQ retransmissions. For similar reasons, at high offered load, we also allow the gNB to use 256QAM 4/5 as the highest MCS, whereas the maximum MCS is limited to 64QAM 9/10 for low load conditions. This is to further reduce the block error probability at low loads,

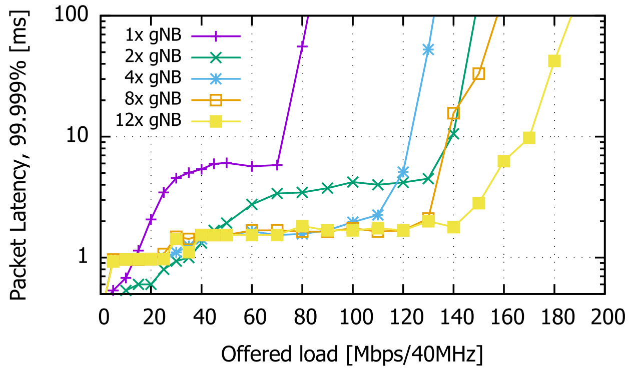


FIGURE 2. NR-FDD: Achievable delay performance versus load and number of installed gNBs.

which is essential in order to meet the 1 ms and 99.999% reliability target. More details on the trade-offs between spectral efficiency and latency are found in [19].

Finally, for the specific case of NR-TDD, a 'DDU' frame structure is adopted, where 'D' and 'U' corresponds to a DL and UL mini-slot of 4 OFDM symbols, respectively. These frequent DL-UL switching points ensure low alignment delay for both the DL data transmission and the HARQ ACK/NACK feedback from the UE. Fast transmission of the UE's HARQ feedback reduces the time between the initial data transmission and eventual retransmissions, thus improving the latency performance.

For NR-FDD, in Fig. 2 we show the achieved DL packet latency at the 99.999 level for different offered loads using the assumptions described in Section II. As the load is increased, generally the achieved packet latency at the 99.999 % level increases. This means that the load needs to be contained in order to guarantee a certain level of performance for a critical IIoT application. For a 1 ms delay critical application, it is preferable to operate with only 1 or 2 gNBs to reduce inter-cell interference and thus the need for HARQ retransmissions. Particularly, a load of up to about 35 Mbps, or an equivalent spectral efficiency of 0.85 bps/Hz, can be supported in the system when 2 gNBs are deployed. However, for delay tolerant applications (e.g. II, III and IV in Table 1), it is possible to add more gNBs to the deployment, as this improves the load that can be served while maintaining a certain latency performance. As an example, for a 4 ms application, 12 gNBs support 155 Mbps compared to 28 Mbps for a single gNB and 95 Mbps for 2 gNBs. The gain of going from 2 to 12 gNBs is however more moderate at ex. 10 ms latency requirements. This aspect would be further improved with directional antennas enabling further scalability with 5G NR.

Results for NR-TDD are shown in Fig. 3. Similar trends as for NR-FDD are observed, although the TDD duplexing mode means that sub-ms latency is slightly more challenging. Therefore, with a low number of gNBs, the offered load while achieving URLLC is limited to around 20 Mbps, or 0.5 bps/Hz, for 2 gNBs. A single gNB serves 1 ms applica-

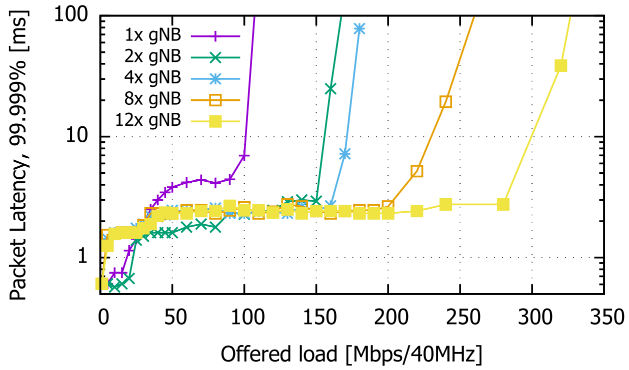


FIGURE 3. NR- TDD: Achievable delay performance versus load and number of installed gNBs.

tions at nearly the same load as 2 gNBs (18 Mbps). Approximately 105 Mbps capacity can be achieved if targeting 10 ms latency target with a single gNB. By going to 2 gNBs, this increases to 155 Mbps and further up to 300 Mbps for 12 gNBs. Results confirm a high level of scalability of NR-TDD, allowing efficient scaling of performance in constrained bandwidth by adding more gNBs.

When comparing NR-FDD with NR-TDD, the former is significantly better for the 1 ms delay requirement as the paired spectrum allows DL and UL transmission simultaneously; whereas for NR-TDD additional delay is introduced as some packets need to wait for the next DL transmission opportunity. However, NR-TDD offers significantly better performance at high offered load regimes, since the TDD access scheme allows asymmetrical DL-UL ratios (in our case, 66 % and 33 % of the 40 MHz bandwidth is used for DL and UL, respectively), instead of the fixed 50 % - 50 % DL-UL ratio of FDD.

B. WI-Fi 6

Wi-Fi 6 uses LBE mode where APs asynchronously perform the mandatory channel access mechanism, i.e. the assessment of the availability of the channel before transmission. The duration of the sensing and the maximum allowed time that a node can occupy the channel (before another sensing procedure must be performed) are determined based on the selected QoS access category [20]. The higher the priority of the data, the lower the duration of the sensing and the maximum transmit opportunity (TxOP) duration. When targeting delay critical traffic, the maximum TxOP duration needs to be limited. To do so, each AP must broadcast the desired TxOP settings to all connected STAs, which means overriding the default TxOP parameters associated with e.g. background and best effort traffic. Based on our simulation results, and operating within regulation guidelines, the optimal maximum TxOP duration should be limited to 1.5 ms and accompanied with short contention windows (CWs) between 3 and 7 clear channel assessment (CCA) slots (one CCA slot is 9 µs). Using shorter CWs, ensures more frequent channel access which

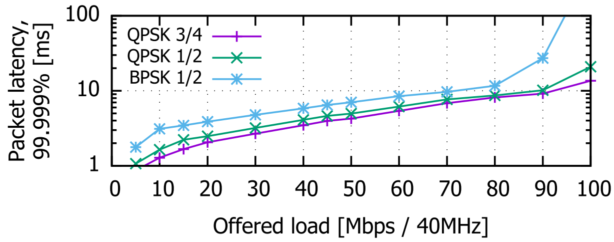


FIGURE 4. Impact of selected BACK MCS on achievable Wi-Fi 6 performance.

reduces potential queuing delay. The drawback of shorter TxOP and CWs is an increase in the overhead and a lower spectral efficiency.

Wi-Fi 6 supports 1024QAM which reduces transmission times for users in good channel conditions and reduce delay of queued users. Further, Wi-Fi 6 supports Orthogonal Frequency Division Multiple Access (OFDMA) for multiple access in a single spatial stream [21]. OFDMA significantly reduces the amount of overhead and channel access time associated with scheduling multiple users simultaneously, especially with smaller data packets. A single trigger frame (TF) is transmitted from the AP to initialize the DL transmission for all the scheduled users and indicates which sub-channel is assigned to which user. Similarly, all scheduled users can transmit their acknowledgements on the sub-channel allocated in the TF. Up to 9 users can be scheduled simultaneously in a 20 MHz channel, sharing the TF and acknowledgment time slots. Without OFDMA, users are typically scheduled one by one consecutively, each one with the associated transmission overhead and channel sensing.

In the used link adaptation method, an inner loop chooses the appropriate transmission MCS based on Rx power (RxP) estimate in the beginning of each TxOP. The Outer Loop Link Adaptation (OLLA) adjusts the initial MCS selection for changes in interference and RxP. OLLA is based on MPDU Acknowledgment (ACK) and implicit Negative Acknowledgement (NACK), where erroneous reception or NACK is determined by missing ACK. As another consideration, a sizeable amount of each TxOP is used to transmit the Block Acknowledgments (BACK) from STAs to the AP in uplink. Selecting the lowest MCS (BPSK, code rate 1/2) for BACK transmission ensures the highest probability of successful decoding, but occupies more time for every single TxOP. In the given scenario with relatively small distance between STAs and serving AP, the minimum MCS is not required for successfully decoding. By increasing the MCS of the BACK transmissions to QPSK 3/4, the packet latency at the 99.999 percentile is improved by releasing more resources for data as illustrated in Fig. 4.

We use a latency optimized multi-user packet scheduler that exploits OFDMA. The maximum number of users are scheduled per TXOP to reduce the amount of overhead per user. As all users are the same QoS class, the users with

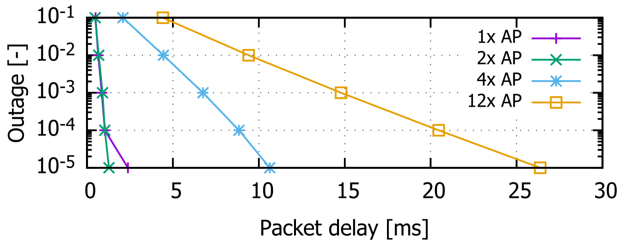


FIGURE 5. Packet latency vs outage for different number of APs (10 Mbps/40MHz load).

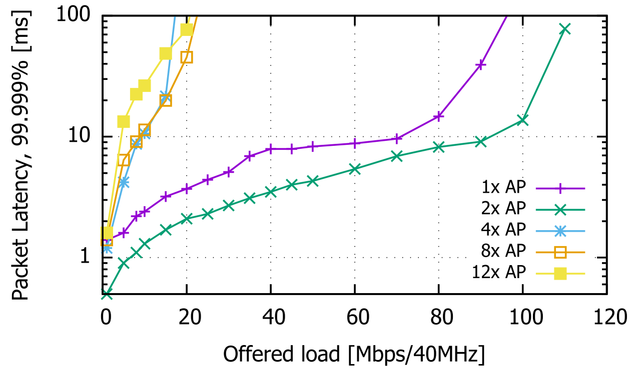


FIGURE 6. Wi-Fi 6: Achievable delay performance versus load and number of installed APs.

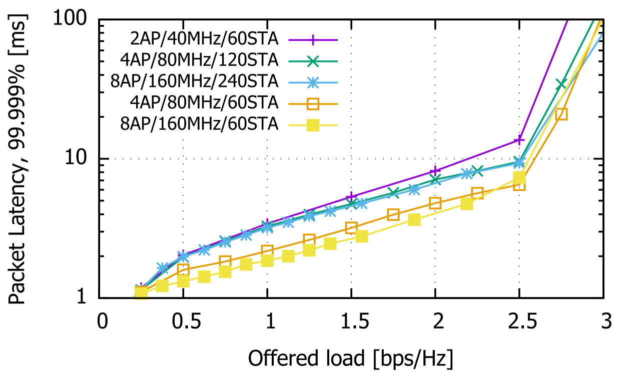


FIGURE 7. Wi-Fi 6: Comparison of delay and load performance versus bandwidths and STA density (20 MHz APs in dedicated spectrum).

the oldest packets are scheduled with a First Come First Served (FCFS) principle. Packet fragmentation is disabled to ensure FCFS functionality for improved delay performance over maximizing spectral efficiency as otherwise any user(s) with fragmented packets takes priority over FCFS scheduling.

Next, the optimal frequency planning for Wi-Fi 6 is explored. The smallest Wi-Fi channel that can be assigned to a single access point (AP) is 20 MHz, and multiple neighbouring channels can be combined into 40, 80 or 160 MHz channels for larger bandwidths. For interference free operation, the considered bandwidth of 40 MHz can be assigned to a single AP as a single 40 MHz channel, or alternatively split evenly between two APs as two 20 MHz channels. For any larger number of APs, interference free operation is not possible with 40 MHz, multiple APs (and associated users) would have to share the same channel(s). Under shared channel, the latency is impacted negatively by multiple APs contenting for channel access and the delay of any collisions. In Fig. 5, four different deployments have been compared at a low network load of 10 Mbps. For each set of simulations, the packet delay statistics have been analyzed and the plot shows what packet latency can be achieved at a certain reliability level (from 99 % up to 99.999 %). In the cases of 4 and 12 APs, the channels are assigned such that APs on the same channel have the largest possible distance between them to minimize cross-interference. Nevertheless, the resulting interference and channel contention in these cases negatively impacts the achievable reliability and delays as shown in Fig. 5. Even when assuming detailed frequency planning for 4 and 12 APs, the interference free cases of 1 and 2 APs perform much better, as illustrated in the figure. The results for other load values indicate similar trends, i.e. that isolating each AP to its own clean channel (or set of channels) is paramount for delay critical applications.

Finally, the performance of Wi-Fi 6 is considered in Fig. 6. For the single AP case with 40 MHz bandwidth, 10 ms performance was achieved up to an offered load of 70 Mbps and 100 ms performance at a load up to around 95 Mbps. While URLLC-level performance is not achieved for the single AP case, 4 ms latency guarantee was achieved up to a load of around 20 Mbps. Queuing delays for 60 users as well as coverage still limit the performance in the single-AP case, and the results show significant performance gains available when deploying 2 APs in separate 20 MHz bands. In this configuration, a 1 ms latency is achieved at around 7 Mbps, or 0.16 bps/Hz, while latency is confined to within 10 ms for a load up to 90 Mbps, i.e. a 30 % capacity increase over the single AP case. The curves also confirms the earlier indication that sharing the same spectrum among two APs in Wi-Fi is detrimental to achieving a good latency performance, even for cases targeting between 10 ms and 100 ms latency.

As Wi-Fi 6 is limited to 2 AP deployment in 40 MHz, we next discuss the scalability of the achieved results. As shown in Fig. 6, dividing equally the spectrum into two AP locations is optimal which means that the results are scalable if installing additional APs in the same factory locations if more than 40MHz bandwidth is available. Moreover, when adding the new APs in the additional locations depicted in Fig. 1, there would also be a gain due to the better coverage. In Fig. 7, we compare the performance of different number of APs when deploying Wi-Fi 6 over 80 MHz or 160 MHz. In all cases, each AP is deployed in its own 20 MHz bandwidth. As example, “4AP/80MHz/120STA” means that there are 4 APs deployed, each in own clean 20 MHz bandwidth, and the user density is 120 STA over the 80 MHz total bandwidth (and 4 APs). Compared to the baseline results for 2×20 MHz APs with 60 STAs, the results scale well when keeping the user density the same, e.g. versus the “4AP/80MHz/120STA” and “8AP/160MHz/240STA”. There is some small <10 % gain going from 2 to 4 APs due to improved coverage gain, and only negligible benefits of 8 APs. However, if amble spectrum is available that the same 60 users can be spread over a larger bandwidth, ex. "4AP/80MHz/60STA" and "8AP/160MHz/60STA" there is a significant improvement in supported load when increasing the number of APs. Doubling the bandwidth leads to 20-60 % spectral efficiency gains for delay requirements up to 7-8 ms.

C. 5G NR-U

In addition to the described optimizations for NR-TDD in Section III-A, further enhancements are now addressed for optimal 5G NR operation in the unlicensed spectrum. Our target is to reduce the impact of the channel access mechanisms on the packet latency. The learnings from Wi-Fi are also applicable to 5G NR-U. Thus, adopting the highest Channel Access Priority Class (CAPC), equivalent to the QoS access category in Wi-Fi 6, and properly distributing the gNBs in the available spectrum improves the latency performance in 5G NR-U deployments with LBE channel access. Aiming at achieving similar performance as in NR-TDD, we explore the added versatility of NR-U to support both LBE and FBE channel access mechanisms.

FBE enables a network design in which nodes are deployed in overlapping channels while minimizing the channel access impact on the system performance. As introduced in Section II-B, FBE defines a synchronous channel access based on a periodic structure, the Fixed Frame Period (FFP), consisting of an occupancy period followed by an idle period. During the idle period, nodes must cease any transmission and assess the channel availability. During the occupancy period, nodes are allowed to transmit given a previously successful channel access. Configuring the gNBs with perfectly overlapping FFP structure ensures 100 % channel access probability at the first channel occupancy acquisition, i.e. the channel access during the sensing period is always successful. However, non-zero channel access failure probability can be observed during the occupancy period.

For a given TDD frame structure, when switching between downlink and uplink or vice versa, nodes must perform the channel access procedure to continue the channel utilization. Adopting a flexible TDD frame selection in each gNB could lead to potential channel access failures. To cope with potential TDD frame selection misalignment and channel access failures, we adopt a fixed TDD frame structure during each gNB occupancy period. More details about the coordinated solution and its performance are documented in [13]. Aligning the downlink-to-uplink and uplink-to-downlink transitions within the TDD frame removes the uncertainty introduced by the channel access procedures. This ensures 100 % success probability in the direction switches. However, non-zero channel access failures might still be observed during the TDD frame. Assuming that the common TDD frame is DDUDDU, there can be cases in which there is no data available to transmit in the uplink-to-downlink transition during the gNB occupancy period. If data becomes available in the next DL TTI, a gNB could be blocked by neighbour gNBs

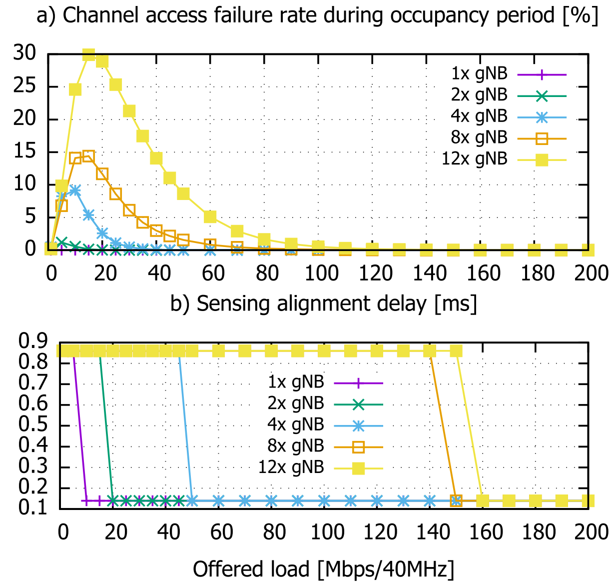


FIGURE 8. NR-U FBE KPI analysis for different network configurations.

with a continuous downlink transmission which was started in the previous DL TTI. The gNBs channel access failure rate during the occupancy time is depicted in Figure 8a. It is noted that the channel access failure rate is dependent on both the load and on the number of gNBs. Specifically, it is highest at medium loads and for high number of gNBs as in this case, the probability of initiating a DL transmission in the middle of a gNB occupancy period and while another gNB is transmitting on the channel is highest. Consequently, the 12 gNBs deployment shows the highest channel access blocking probability at an offered load of 15 Mbps.

It is also worth noticing that an additional delay contribution is introduced with FBE operation. Nodes are only allowed to access the channel during the sensing period, and therefore, a certain time is dedicated on waiting for the next channel access opportunity. This is denoted as sensing alignment delay and it is dependant on the FFP duration. In order to minimize the sensing alignment delay, the minimum FFP duration among those supported by 5G NR-U (i.e. 1 ms) is recommended. This ensures that the sensing alignment delay is upper bounded to less than 1 ms. The main cost of setting the minimum FFP duration is a decrease in the channel utilization as more sensing intervals are introduced. The sensing alignment delay is dependant on the offered traffic per gNB. When the load per gNB is high, data packets always being available for scheduling at the beginning of every FFP ensures that the sensing alignment delay cannot exceed the duration of the idle period, i.e. 4 OFDM symbols in our assumptions. On the other hand, when the load per gNB is low, the sensing alignment delay can be as high as 24 OFDM symbols according to our assumptions. This can be observed in Figure 8b in which the sensing alignment delay at 99.999 % reliability is shown over a range of offered loads for the considered deployments. It can be noted that the alignment delay takes 2 values at the given reliability: i) 4 OFDM symbols if data is available for transmission at the

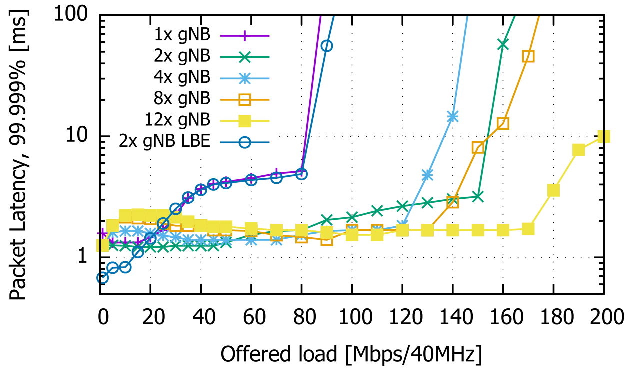


FIGURE 9. **NR-U: Achievable delay performance versus load and number of installed gNBs. FBE mode except where indicated.**

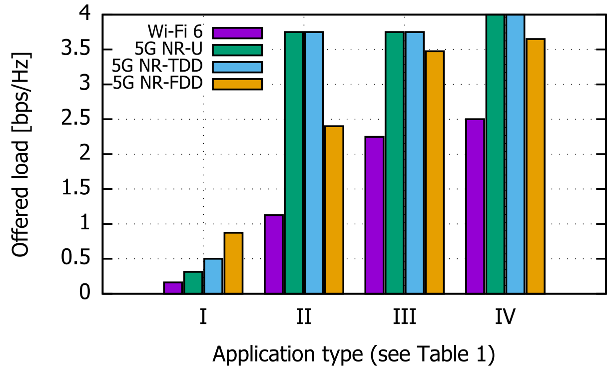


FIGURE 10. **Comparison of offered load for different applications (2 AP/gNBs deployed, NR-U uses LBE for type I, FBE for types II, III, and IV).**

beginning of the FFP or ii) 24 OFDM symbols if gNBs need to wait for a full occupancy period to be finished before sensing. These values represent the best-case and worst-case of the sensing alignment delay respectively. Any of the considered scenarios shows the maximum sensing alignment delay for the low end of the offered load. The transition to the minimum value depends on the offered load per gNB. A higher system offered load is needed for denser deployments to achieve the minimum sensing alignment delay. For instance, 50 Mbps and 160 Mbps system loads are needed to achieve minimum sensing alignment delay for the 4 gNBs and 12 gNBs scenarios, respectively. The addition of the sensing alignment into the packet delay budget implies that a constant delay requirement of 1 ms at 99.999 % reliability cannot be met consistently. This very stringent latency-reliability target could only be met under optimal conditions of queuing delay. On the other hand, without data being queued, the sensing alignment at 99.999 % reliability shows its worst-case value. This prevents the support of Type I reference application with FBE. For such applications, LBE channel access operation is preferred since it removes the sensing alignment delay. The adoption of LBE is exclusively reserved for scenarios with reduced number of gNBs per channel and low offered load.

In Fig. 9 the packet latency at 99.999 % reliability for NR-U FBE are depicted. Additionally, the performance for NR-U LBE with 2 gNBs is included. Focusing first on the FBE performance, it is noted that an offered load up to approximately 80 Mbps is supported by the single gNB scenario before the packet delay increases abruptly due to very high queuing delays. Adding an extra gNB to the scenario alleviates significantly the queuing issues as well as it improves the UEs coverage. As a result, the capacity nearly doubles as compared to the single gNB case. The supported load for 2 gNBs in FBE is close to 150 Mbps when the delay requirement is relaxed to 4 ms or higher. With 4 gNBs, the benefits of a better distribution of the load across a higher number of gNBs are overtaken by the corresponding decrease in experienced SINR, resulting in an increased resource utilization per packet and an overall performance decrease as compared to 2 gNBs. A similar effect is shown for the 8 gNBs scenario. For 12 gNBs, the benefits from reduced load per gNB are finally visible, with 20 % capacity increase as compared to 2 gNBs as loads up to 200 Mbps can be supported while guaranteeing a packet latency below 10 ms. Adding 6 gNBs only shows limited returns compared to the additional infrastructure complexity. Scenarios with 4, 8 and 12 gNBs show similar trend at low loads. The packet delay initially increases and it becomes lower as more offered load is supported. The cause of this behaviour are the effects described in Figure 8.

Generally, the NR-U FBE deployments are challenged to deliver URLLC services with strict 1 ms packet latency. As shown in Fig. 9, for very stringent latency-reliability requirements, LBE channel access must be adopted. The 2 gNB scenario is capable of supporting 12.5 Mbps, or 0.31 bps/Hz, within the delay constraints. The results also show the value of FBE over LBE by comparing the 2 gNB cases. Due to better spectrum utilization, FBE operation leads to more than 3× higher offered load than LBE at the 4 ms target latency and nearly 2× at 10 ms and beyond.

IV. PERFORMANCE COMPARISON

After providing design guidelines for the different radio technologies to maximize their achievable capacity under a certain latency target, we conduct in this section a full-scale comparison. The comparison metric is the maximum offered load (capacity), i.e. the maximum amount of traffic that can be served by the deployment, without violating the QoS requirement. The technologies are compared in Figure 10 for a deployment of 2 APs/gNBs respectively as this was found to be a good reference deployment for all technologies. In this section, we simplify the notation to bps/Hz, to allow easier assessment of how much spectrum is needed to achieve a certain offered load. Scalability of results were discussed in previous sections.

Wi-Fi 6 can support reliable sub-ms performance in this fully controlled environment at a load <0.16 bps/Hz. As discussed, NR-U only supports reliable sub-ms performance when adopting LBE mode of operation, i.e. the same channel access scheme as Wi-Fi 6 and those results are used for the Application Type I. At a supported load of 0.31 bps/Hz, NR-U shows a 90 % gain over Wi-Fi 6. Using FBE mode of operation when the delay requirement is relaxed to >1-2 ms, NR-U achieves the same performance as licensed 5G technologies.

NR-TDD achieves an offered load of 0.5 bps/Hz for Application I. It is a 3-fold performance increase as compared to Wi-Fi and 1.6× better than NR-U. NR-TDD benefits from a larger bandwidth availability per deployed gNB/AP, 40 MHz/gNB for NR-TDD vs 20 MHz/gNB for Wi-Fi and NR-U with LBE. For Type I applications, 5G NR-FDD clearly outperforms other technologies due to its ability for simultaneous UL and DL transmission, e.g. approximately 0.72 bps/Hz which is >4× better than Wi-Fi 6 and 2.8× better than NR-U. If restricted to a single gNB/AP installed (not shown in Figure 10), NR-TDD however outperforms other solutions' URLLC performance with supported 0.45 bps/Hz load. As application requirements are relaxed, Wi-Fi 6 starts to support significantly higher load, ex. 1.125 bps/Hz at the 4 ms delay requirement (II) and around 2.25-2.5 bps/Hz for the 10-100 ms requirement (III and IV). When relaxing the delay requirement to slightly above 1 ms, both NR-U and NR-TDD shows very high performance as the fundamental delays caused by frame alignment and TDD uplink/downlink switching delays become secondary and benefits of high frequency reuse, more advanced link adaptation and channel measuring techniques appear. At the 4 ms delay requirement, the performance of NR-U, NR-TDD, and NR-FDD is 2-3.3× better than Wi-Fi 6 at around 3.5 bps/Hz. The benefit over Wi-Fi is reduced to 1.4-1.6× when relaxing the delay requirement to 10-100 ms.

As a limitation of the study, mobility has not been explicitly modeled in the simulations although it is assumed that a device is always connected to the best gNB/AP. In NR and NR-U, maintaining URLLC performance during handovers is verified for some forms of mobility [22], but required use of ex. multiple transmit and receive points (TRP) or the Dual Active Protocol Stack (DAPS) [1]. Although Wi-Fi 6 may also apply handover optimizations such as IEEE 802.11r, 802.11k, and 802.11v extensions, fully seamless handover is not generally achievable. Some proprietary IIoT optimized Wi-Fi 6 solutions indicates that handover performance latency can be as low as a few or a few tens of ms, but reliability levels are not reported.

V. CONCLUSION

This paper reports on the capabilities of Wi-Fi 6, 5G unlicensed (NR-U), and 5G licensed (NR) technologies in terms of achieving guaranteed low packet latency at high reliability levels in a downlink oriented IIoT scenario. The paper also presents optimization and deployment guidelines for each radio technology to serve critical IIoT traffic while maximizing the available capacity. The reported performance levels should be understood as the performance potential that the wireless technology supports if the system operator and vendor focus fully on reliable latency minimization.

For Wi-Fi 6, operation in a fully controlled environment with no interference in the same spectrum carrier is critical for an AP to deliver consistent low latency performance. Hence, frequency planning and interference control are essential mechanisms as part of Wi-Fi IIoT deployment. Additionally, we presented a scheduler optimization strategy that ensures low latency performance at the expense of higher channel overhead and thus lower spectral efficiency. When optimized and operating in a low load up to 0.16 bps/Hz, Wi-Fi 6 can achieve ultra-reliable low latency performance (i.e. <1 ms packet latency at 99.999 % reliability). Significantly higher offered load can be carried when the target latency is relaxed but still guaranteed, from 1.13 bps/Hz with a 4 ms requirement up to 2.5 bps/Hz for a 100 ms requirement.

For the FBE variant of NR-U, the offered load for application latency requirements above 2 ms is very close to that of 5G licensed technologies at around 3.75 bps/Hz, and 2-3× better than using LBE operation. LBE mode is however needed to achieve strict URLLC performance, and can support 0.31 bps/Hz or nearly 2× that of Wi-Fi 6 due to more elaborate link adaptation mechanisms. As for Wi-Fi 6, non-system interference must be tightly controlled to guarantee tight latency performance.

5G NR-FDD has superior URLLC performance and meets the sub-ms delay requirement at >5× higher load than Wi-Fi 6 and >2× higher than NR-U LBE. At the 4-100 ms delay targets, NR-FDD supports around 2.4-3.6 bps/Hz. NR-TDD shows very similar performance to NR-FDD and NR-U due to the larger downlink transmission bandwidth, but offers lower performance than NR-FDD for 1 ms delay targets due to the TDD uplink/downlink switching delays.

When more capacity is needed than achievable in the considered 40 MHz bandwidth, Wi-Fi 6 needs to deploy more clean spectrum and it was shown how the results can be scaled to wider bandwidth. The benefit of especially NR-TDD, and to some extent NR-FDD and NR-U, is that capacity can be increased while maintaining the latency target by deploying more gNBs in the same spectrum. This is useful especially for licensed deployments where spectrum is often more scarce.

REFERENCES

[1] *NR; NR and NG-RAN Overall Description; Stage-2 (Release 16)*, document TR 38.300 V16.4.0, 3GPP, Dec. 2020.

[2] T.-K. Le, U. Salim, and F. Kaltenberger, "An overview of physical layer design for ultra-reliable low-latency communications in 3GPP releases 15, 16, and 17," *IEEE Access*, vol. 9, pp. 433–444, 2021.

[3] D. Weller, R. D. Mensenkamp, A. V. D. Vegt, J.-W.-V. Bloem, and C. D. Laat, "Wi-Fi 6 performance measurements of 1024-QAM and DL OFDMA," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–7.

[4] G. Cena, S. Scanzio, and A. Valenzano, "Seamless link-level redundancy to improve reliability of industrial Wi-Fi networks," *IEEE Trans. Ind. Informat.*, vol. 12, no. 2, pp. 608–620, Apr. 2016.

[5] M. Mhedhbi, M. Morcos, A. Galindo-Serrano, and S. E. Elayoubi, "Performance evaluation of 5G radio configurations for industry 4.0," in *Proc. Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Oct. 2019, pp. 1–6.

[6] M. Iwabuchi, A. Benjebbour, Y. Kishiyama, G. Ren, C. Tang, T. Tian, L. Gu, T. Takada, and T. Kashima, "5G field experimental trials on URLLC using new frame structure," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2017, pp. 1–6.

[7] R. Maldonado, C. Rosa, and K. I. Pedersen, "Analysis of high-reliable and low-latency communication enablers for new radio unlicensed," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, May 2020, pp. 1–6.

[8] R. B. Abreu, G. Pocovi, T. H. Jacobsen, M. Centenaro, K. I. Pedersen, and T. E. Kolding, "Scheduling enhancements and performance evaluation of downlink 5G time-sensitive communications," *IEEE Access*, vol. 8, pp. 128106–128115, 2020.

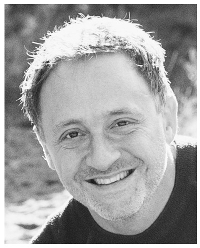
[9] M. Hirzallah, M. Krunz, B. Kecicioglu, and B. Hamzeh, "5G new radio unlicensed: Challenges and evaluation," *IEEE Trans. Cognit. Commun. Netw.*, early access, Dec. 2, 2020, doi: 10.1109/TCCN.2020.3041851.
[10] *Study on Channel Model for Frequencies From 0.5 to 100 GHz (Release 16)*, document TR 38.901 V16.1.0, 3GPP, Nov. 2020.
[11] *Feasibility Study on Licensed-Assisted Access to Unlicensed Spectrum (Release 13)*, document TR 36.889 V13.0.0, 3GPP, Jul. 2015.
[12] *5 GHz RLAN; Harmonised Standard Covering the Essential Requirements of Article 3.2 of Directive 2014/53/EU*, Standard ETSI EN 301 893 V2.1.1, May 2017.
[13] R. Maldonado, C. Rosa, and K. I. Pedersen, "A fully coordinated new radio-unlicensed system for ultra-reliable low-latency applications," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, May 2020, pp. 1–6.
[14] R. F. W. Coates, G. J. Janacek, and K. V. Lever, "Monte Carlo simulation and random number generation," *IEEE J. Sel. Areas Commun.*, vol. JSAC-6, no. 1, pp. 58–66, Jan. 1988.
[15] *Indoor Industrial Channel Model Calibration Results*, document TR R1-1909704, 3GPP, May 2019.
[16] G. Pocovi, K. I. Pedersen, and P. Mogensen, "Joint link adaptation and scheduling for 5G ultra-reliable low-latency communications," *IEEE Access*, vol. 6, pp. 28912–28922, 2018.
[17] *NR; Physical Layer Procedures for Data (Release 16)*, document TR 38.214 V16.4.0, 3GPP, Jan. 2021.
[18] *Discussion on the RAN2 LS on TSN Requirements Evaluation*, document TR R1-1813120, 3GPP, Nov. 2018.
[19] G. Pocovi, B. Soret, K. I. Pedersen, and P. Mogensen, "MAC layer enhancements for ultra-reliable low-latency communications in cellular networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2017, pp. 1005–1010.
[20] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements*, Standard IEEE 802.11e IEEE 802.11 WG, 2015
[21] E. Khorov, A. Kiryanov, A. Lyakhov, and G. Bianchi, "A tutorial on IEEE 802.11ax high efficiency WLANs," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 197–216, 1st Quart., 2019.
[22] *Study on Self Evaluation Towards IMT-2020 Submission (Release 16)*, document TR 37.910 V16.1.0, 3GPP, Oct. 2019.

ROBERTO MALDONADO received the B.Sc. and M.Sc. degrees in telecommunication engineering from the University of Granada, Spain, in 2014 and 2016, respectively, and the Ph.D. degree from Aalborg University, Denmark, in 2021. He is currently working as a 5G Researcher with Nokia Bell Labs. His research interests include ultra-reliable and low-latency communications, 5G and beyond radio resource management, and unlicensed spectrum technologies.

ANDERS KARSTENSEN received the M.Sc. and Ph.D. degrees from Aalborg University, Denmark, in 2015 and 2020, respectively. He currently holds a postdoctoral position at Aalborg University. His research interests include Wi-Fi and 5G Industrial IoT and spectrum sharing.

GUILLERMO POCOVI received the M.Sc. degree in telecommunications engineering from the Universitat Politècnica de Catalunya, in 2014, and the Ph.D. degree from Aalborg University, Denmark, in 2017. He is currently with Nokia Bell Labs, Aalborg. His research interests include the standardization of ultra-reliable and low-latency communications (URLLC) and the Industrial Internet of Things (IIoT) use cases in 5G new radio.

ALI A. ESSWIE received the M.Sc. degree from Memorial University, Canada, in 2017, and the Ph.D. degree from Aalborg University, in 2020. The present work was done while he was with Nokia Bell Labs, Aalborg, as a Standardization Research Engineer. From 2013 to 2016, he acted as a Wireless Standards Engineer with the Intel Labs and Huawei Technologies. He has authored more than 27 publications and 12 patent filings. His main research interests include 5G new radio, wireless ML/AI, radio resource management, ultra-reliable and low latency communications, massive MIMO, and power saving.



CLAUDIO ROSA received the M.Sc.E.E. degree from Aalborg University, in 2000, the M.Sc.E.E. degree in telecommunication engineering from the Politecnico di Milano, Italy, in 2003, and the Ph.D. degree from Aalborg University, in 2005. He joined Nokia in 2005, where he contributed to standardization of 4G and 5G systems working on uplink power control and general radio resource management, carrier aggregation, dual connectivity, and unlicensed spectrum operation. He holds more than 30 granted U.S. patents and has coauthored more than 50 scientific publications. His current research interests include dynamic spectrum sharing and private network deployments.

OLLI ALANEN received the M.Sc. and Ph.D. degrees from Jyväskylä University, Finland, in 2004 and 2007, respectively. He is the author of more than 30 academic publications and more than 90 patent applications. Recently, he has been interested on reliable low latency Wi-Fi and 6G simulation platforms. His research interests include Wi-Max, LTE, and Wi-Fi radio communication related topics with focus on simulations, standardization, and inventing.

MIKA KASSLIN (Member, IEEE) is currently the Department Head with Nokia Bell Labs, where he focuses on wireless technologies deployed at and designed for unlicensed bands. Since 1997, he has been conducting research on Wi-Fi with a focus on techniques related to *ad hoc* networking and mobile terminal power save. He has contributed to many 802.11 standard versions and more than 200 patent applications.

TROELS KOLDING received the M.Sc. and Ph.D. degrees from Aalborg University, Denmark, in 1996 and 2000, respectively. His M.Sc. degree was achieved in collaboration with the Wireless Information Network Laboratory (WINLAB), NJ, USA. Since joining Nokia in 2001, he has been active in research and management for standardization, network architecture, and portfolio management. He holds more than 50 granted U.S. patents and the author of more than 80 scientific publications. His current research interests include 5G IIoT, time-sensitive communications, time-synchronization, and 5G/6G radio resource management and spectrum sharing.